

Conserving Alfalfa Wild Relatives: Is Past Introgression with Russian Varieties Evident Today?

Stephanie L. Greene,* Theodore J. Kisha, and Nikolai I. Dzyubenko

ABSTRACT

Central Asia, particularly Kazakhstan, supports a rich concentration of wild alfalfa (*Medicago sativa* L.) relatives. Because tetraploid wild alfalfa freely crosses with domesticated alfalfa, they are important genetic resources. When identifying in situ populations to conserve, contamination of wild populations with domesticated alleles is an important consideration. We evaluated population structure and introgression between six wild populations of *M. sativa* nothosp. *varia* collected in northwestern Kazakhstan and five traditional Russian cultivars historically grown in the same region using two amplified fragment length polymorphism primer pairs and six simple sequence repeat loci. We found no difference between the Russian cultivars and wild populations for number of alleles or percentage polymorphic loci; however, gene diversity was less in the wild than in the cultivated populations. Cluster analysis and principle component analysis showed clear separation between wild and cultivated populations. Genetic differentiation among the cultivars was less than among the wild populations. Using a Bayesian approach, we found limited evidence of admixture among the wild and cultivated forms, although more admixture was evident in wild populations collected in less-remote areas. On the basis of marker data, we concluded that three of the six wild populations stood out as candidates for in situ conservation given their uniqueness and lack of admixture with cultivated forms.

S.L. Greene, USDA-ARS, National Temperate Forage Legume Germplasm Resources Unit, 24106 N. Bunn Rd., Prosser, WA 99350; T. Kisha, USDA-ARS, Western Regional Plant Introduction Station, Johnson Hall, Room 59, Washington State Univ., Pullman, WA 99164; N. Dzyubenko, N.I. Vavilov Institute, 42, Bolshaya Morskaya Str., St. Petersburg, Russia, 196600. Received 5 Dec. 2007. *Corresponding author (stephanie.greene@ars.usda.gov).

Abbreviations: AFLP, amplified fragment length polymorphism; PCA, principle component analysis; PI, plant inventory; SSR, simple sequence repeat.

THE EVOLUTIONARY CONSEQUENCES of gene flow between domesticated crops and their wild relatives are widely recognized (Ellstrand et al., 1999; Jarvis and Hodgkin, 1999; Ellstrand, 2003; Lefèvre, 2004). From a conservation perspective, the consequences are frequently negative for wild relative species. Domestic alleles have been shown to impact fitness in wild populations (Ellstrand, 2003; Hails and Morley, 2005). Hoc et al. (2006) reported degeneration of hybrid progeny between *Phaseolus vulgaris* var. *aborigineus* and domesticated beans growing in Argentina. Introgression between wild and cultivated forms has also resulted in the development of highly aggressive weeds (Panetsos and Baker, 1967; de Wet and Harlan, 1975; Boudry et al., 1993). Hybridization with cultivated crops can also threaten wild relatives through outbreeding depression or genetic assimilation (Hails and Morley, 2005). For example, *Oryza rufipogon* ssp. *formosana* is considered nearly extinct due to hybridization with *O. sativa* (Oka 1992), and genetic erosion is predicted for common wild rice (*O. rufipogon*) throughout China due to hybridization with domesticated rice (Song et al., 2005).

Published in Crop Sci. 48:1853–1864 (2008).

doi: 10.2135/cropsci2007.12.0668

© Crop Science Society of America

677 S. Segoe Rd., Madison, WI 53711 USA

All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher.

Alfalfa (*Medicago sativa* ssp. *sativa* L.), an important animal forage, makes a significant but often overlooked contribution to agriculture. Wild relatives in the primary alfalfa gene pool consist of an interfertile complex of autotetraploid ($4n = 32$) subtaxa within *M. sativa* L. (taxonomic nomenclature after Wiersema et al., 1990). Historically, alfalfa wild relatives have contributed substantially to the development of the highly productive modern varieties seen today (Sinskaya, 1961; Lesins and Lesins, 1979; Michaud et al., 1988). As a consequence of the sympatric occurrence of domestic alfalfa and alfalfa wild relatives and the propensity of tetraploids to intercross, one would expect wild populations to be impacted by gene flow from domestic sources. In Switzerland, Rufener Al Mazyad and Ammann (1999) reported that wild populations of tetraploid *M. sativa* ssp. *falcata* are endangered and have disappeared from valleys where it was recorded as growing in the 1960s. They attribute the disappearance to the use of cultivated alfalfa and subsequent development of hybrid swarms between the wild tetraploid *M. sativa* ssp. *falcata* and domesticated alfalfa. In Spain, there is evidence that gene flow occurs between cultivated alfalfa and a local wild form of *M. sativa* ssp. *sativa*, 'Mielga', distinguished by a prostrate growth habit and rhizomes (Jenczewski et al., 1999a,b; Muller et al., 2001, 2003). These studies used quantitative trait, allozyme, randomly amplified polymorphic DNA, and mitochondrial markers to show that three types of "wild" populations existed: those showing no evidence of gene flow, those intermediate between domesticated and wild based on quantitative and neutral markers, and those populations that had the characteristic wild phenotype but were similar to the domesticated types based on neutral markers. The latter type was found in southern Spain, where cultivated alfalfa is irrigated, and the wild-type growth habit may be more adapted to the drier climate.

Central Asia, including the Republic of Kazakhstan, is considered to be an important center of diversity for alfalfa. Ivanov (1988) and colleagues studied the distribution of intra- and interspecific diversity and hybridization in Kazakhstan and advocated five specific geographic areas be protected. In situ reserves were not designated at that time, nor since (Meilleur and Hodgkin, 2004). In 2000, a joint U.S.–Russian–Kazakhstan seed-collecting trip was made in northwestern Kazakhstan, and 135 accessions of wild alfalfa relatives were collected (Greene et al., 2005). The collecting route went through the Mugodzharo–Ustyurt center of introgressive hybridization proposed by Ivanov (1988). The collectors surveyed and collected from naturally occurring populations at locations proposed as in situ reserves for alfalfa relatives by Ivanov (1988). Although alfalfa was no longer cultivated in the region in 2000, it had been cultivated in the past (Ivanov, 1988). If Russian cultivars had been introduced into Kazakhstan,

one would expect that intensive gene flow had taken place between the domesticated and indigenous wild forms. Recognizing the negative consequences of introgression between domestic and wild species, we conducted this study to examine the consequences of gene flow between Russian alfalfa varieties historically grown in the area and local indigenous populations of alfalfa wild relatives. The results of this study could support decisions regarding the identification of wild relative populations least impacted by past introgression with domestic alleles.

Given the genetic complexity of the populations studied (i.e., cultivated and wild populations of autotetraploid, outcrossing perennial species), we felt an analysis using multiple marker systems would be more informative. Simple sequence repeat (SSR) polymorphisms or microsatellite markers are multiallelic and codominant and thus highly informative at a single locus (Mariette et al., 2002). Because of their codominant nature, they are used widely in studies on population genetic structure and are generally considered effective for understanding fine-scale structure within populations (Balloux and Lugon-Moulin, 2002; Gaudeul et al., 2004; Nybom, 2004). Ellwood et al. (2006) found that relatively small numbers of loci can be effective in resolving questions of relatedness. Although it is difficult to score exact allele dosage in polyploidy species, estimates have been reported in autotetraploid species using codominant data (Nybom, 2004; Flajoulot et al., 2005). However, SSRs are more costly to develop and run, null alleles can be generated due to variability in primer sites, and mutation patterns appear to be highly complex, making it difficult to determine the appropriate mutation model and subsequent statistics to infer genetic structure (Balloux and Lugon-Moulin, 2002; Mariette et al., 2002).

In contrast to SSR markers, amplified fragment length polymorphism (AFLP) markers are biallelic and dominant. Although less informative of genetic structure, they allow for the efficient sampling of many loci (Powell et al., 1996; Gaudeul et al., 2004). Thus, AFLPs lend themselves to studies in which more loci are needed to estimate diversity because genomic heterogeneity is high (Mariette et al., 2002). Despite being dominant markers, AFLPs have shown themselves effective in discriminating among populations, and correctly assigning individuals to populations, compared with SSRs (Gaudeul et al., 2004; Woodhead et al., 2005).

Mariette et al. (2002), found that in studies of wild plant populations, estimates for within-population diversity made with different markers were not generally correlated. In a series of simulations, they found that low correlations occurred (i) when populations were not different enough, (ii) when too few markers were used to sufficiently sample high heterogeneity, and iii) when populations were not in genetic equilibrium (Mariette et al., 2002). In contrast, Nybom (2004) found that estimates

of genetic distance made by SSRs and dominant markers were positively correlated. Estimates of within-population variation were generally greater and ranged wider when SSRs were used, but estimates of among population variation were the same as with dominant markers (Gaudeul et al., 2004; Nybom, 2004; Woodhead et al., 2005). Because the alfalfa taxa we were examining were outcrossing, autotetraploids species, we would expect noncongruence among marker types, since genomic heterogeneity is high.

Our objectives were to use AFLP and SSR markers (i) to investigate the population genetic structure of six wild populations of *M. sativa* nothosp. *varia* collected in northwestern Kazakhstan and five Russian cultivars historically grown in the same region, and (ii) to determine if gene contamination occurred between the wild and cultivated species. Although our overall aim was to examine the impact domesticated Russian varieties may have had on populations of local wild relatives, in an effort to help identify wild relative populations that merit in situ protection, a secondary aim was to compare the two types of markers to see if they gave similar results in an autotetraploid species.

MATERIALS AND METHODS

Plant Material

Eleven accessions, representing six wild populations of *M. sativa* ssp. *varia* and five Russian cultivars of *M. sativa* ssp. *sativa* were evaluated (Table 1). Sixteen plants from each of the wild accessions and 14 to 16 plants from the five Russian cultivars were analyzed. The six wild accessions were collected in 2000, during a joint U.S.–Russian–Kazakhstan seed-collecting trip performed in the Aktobe region of northwestern Kazakhstan (Greene et al., 2005). Bulk original seed collected from each

location was used in the analysis. Approximately equal quantities of seed had been collected from 50 to 100 individual plants at each site. PI 634170 and PI 634169 were collected 8 km apart, in a remote, unpopulated area of the southern Mūgodžor Mountains, approximately 275 km southeast from the city of Aktobe. PI 634153 and PI 634154 were collected less than 1 km apart, from a lightly populated area of the northern grassland steppe, 150 km northwest of Aktobe. PI 634145 and PI 634142 were collected in a moderately populated area in the northern grassland steppe, 50 and 75 km southwest of Aktobe, respectively. Following the collection trip, taxonomy was verified through chromosome counts and by examining pod characteristics of original plants and flower color of progeny regenerated in Prosser, WA. The five Russian cultivars had been donated by the N.I. Vavilov Institute to the USDA-ARS National Plant Germplasm System germplasm collection between 1959 and 1986. Three of the five cultivars had passport data that indicated the seed had been collected in the Aktobe region. Historically, all five cultivars were grown in the region (N.I. Dzubenko, personal communication).

AFLP and Microsatellite Markers

Several unexpanded trifoliate buds were removed from each plant and placed in a 1.5-mL microcentrifuge tube. DNA was extracted using the MagneSil Kit from Promega (Madison, WI). Amplified fragment length polymorphism markers were generated using AFLP Analysis System I kits supplied by Life Technologies (Carlsbad, CA) according to their recommended procedures. The selective amplification was modified as a 10-μL reaction, including 0.25 units of Taq polymerase (Hoffmann-La Roche, Basel, Switzerland) and the accompanying buffer, 2 μL of MseI primer from the Life Technologies kit, 0.5 μM of fluorescent-labeled *Eco*RI primer and 2 μL of preamplified DNA diluted 10:1 with 0.1X Tris buffer. Separation and visualization of the markers was performed on 6.5% polyacrylamide using a Li-Cor GeneReadIR Automated AFLP Apparatus (Lincoln, NE).

Table 1. Alfalfa accessions from the USDA-ARS National Plant Germplasm System Collection (NPGS).

Accession no.	<i>Medicago sativa</i>	Improvement status [†]	Location	Description [‡]
PI 634153	nothosp. <i>varia</i>	Wild	50°17'55'' N, 56°5'20'' E	
PI 634170	nothosp. <i>varia</i>	Wild	48°54'04'' N, 58°34'25'' E	
PI 634169	nothosp. <i>varia</i>	Wild	48°49'31'' N, 58°32'08'' E	
PI 634145	nothosp. <i>varia</i>	Wild	50°9'3'' N, 56°58'35'' E	
PI 634154	nothosp. <i>varia</i>	Wild	50°17'55'' N, 56°05'20'' E	
PI 634142	nothosp. <i>varia</i>	Wild	49°54'40'' N, 56°58'35'' E	
PI 430552	ssp. <i>sativa</i> 'Dikorastuskaya'	Cultivar		Cultivar collected from Aktubinsk region, Kazakhstan by VIR; donated to NPGS in 1978 [‡]
PI 258836	ssp. <i>sativa</i> 'Priaralskaya'	Cultivar		Cultivar collected from Aktubinsk region, Kazakhstan by VIR; donated to NPGS 1959 [‡]
PI 258825	ssp. <i>sativa</i> 'Dolanskaya 2'	Cultivar		Cultivar collected from Karaganda region, Kazakhstan by VIR; donated to NPGS 1959
PI 505860	ssp. <i>sativa</i> 'Semirechenskaya'	Cultivar		Local cultivar released in 1934 for Kazakhstan; donated to NPGS by VIR in 1986
PI 258838	ssp. <i>sativa</i> 'Tibetskaya'	Cultivar		Cultivar collected from Aktubinsk region by VIR; donated to NPGS in 1959.

[†]From USDA-ARS Germplasm Resources Information Network (<http://www.ars-grin.gov>).

[‡]VIR, N.I. Vavilov All-Russian Institute of Plant Industry.

Table 2. Microsatellite primers used to study the genetic relationship between six populations of wild alfalfa relatives and five varieties of Russian alfalfa.

Name [†]		Sequence
1. Act001	Fwd [‡]	tac aca agc aat tca agg aag g
	Rev	cac acg act att gcg ctt atg
2. Act003	Fwd	taa ctt cca ttc ttc caa cct g
	Rev	ttc tac atc tgc tct ctg ttg aat c
3. Act005	Fwd	caa tcc gtg agt ggt gag aa
	Rev	ttg gac cga act ggg taa ac
4. Act007	Fwd	ctt ccc ctt cgt ttt tct cc
	Rev	gat gca aac atg tgc cag ac
5. Fca16	Fwd	ggg cga acc aag cat gt
	Rev	taa aaa aca tta cat gac ctc aaa
6. AW31	Fwd	tgt aaa acg acg gcc agt
	Rev	gtg aag act ttg cgg tgg at

[†]1–5, Diwan, et al. (1997); 6. Mary Sledge, Noble Foundation (personal communication, 2006).

[‡]Fwd, forward; Rev, reverse.

A total of six microsatellite primer pair loci were examined (Table 2). Reaction volume was 10 µl and contained 0.1 unit of Biolase Taq polymerase from Bioline (Taunton, MA); 25 ng of template DNA, and reagent concentrations of 150 µM each dNTP, 1.5 mM MgCl₂, and 0.2 µM each of forward and reverse primers. The amplification method started with an initial denaturing step at 94°C for 30 s, followed by 15 cycles beginning at 94°C/10 s, 65°C/30 s, and 72°C/30 s and stepping down the annealing temperature one degree each cycle to 50°C, and ending with 10 cycles of 94°C/10 s, 50°C/30 s, and 72°C/30 s. Separation and visualization of the markers were performed on 6.5% polyacrylamide using a Li-Cor GeneReadIR Automated AFLP Apparatus. Allele dosage was determined from peak heights visualized using GeneImager Software (Scanalytics, Rockville, MD).

Statistical Analysis

Genetic Diversity

Departure from Hardy–Weinberg equilibrium was tested using AUTOTET (Thrall and Young, 2000). Bonferroni corrections were applied over loci analyzed, adjusting *p* values to 0.05. To check for the presence of null alleles, the frequency of punitive null alleles was calculated as $r = (H_e - H_o) / (1 + H_o)$ after Brookfield (1996), where H_e is expected heterozygosity and H_o is observed heterozygosity. For the AFLP markers, each band was treated as a locus and scored as present (1) or absent (0).

For the SSRs, genetic diversity was estimated within each population based on the number of alleles (*A*) at each loci, and H_e , expected heterozygosity, calculated as

$$H_e = \left(1 - \sum p_i^2\right) / (4n - 1)$$

where p_i is the frequency of the *i*th allele and *n* is the number of individuals in the sample, using AUTOTET (Thrall and Young, 2000). Because Julier et al. (2003) reported that double reduction is rare in alfalfa, we estimated H_e assuming random assortment of homologous chromosomes into gametes. For the AFLP markers, we calculated Nei's gene diversity (Nei, 1973)

and percentage polymorphic loci using POPGENE 1.31 (Yeh et al., 1999).

Genetic Distance and Differentiation

We calculated Prevosti's distance,

$$D_p = \frac{1}{l} \sum_{k=1}^l \frac{1}{2} \sum_{i=1}^a |p_{ik} - q_{ik}|$$

where *l* is the number of loci examined, and p_{ik} and q_{ik} are the frequencies of the *i*th allele at the *k*th locus in populations *P* and *Q*, respectively (Prevosti et al., 1975). Prevosti's distance was used to characterize differences in allele frequencies among populations because we were comparing both natural and synthetic populations, and also because AFLP distance is a marker distance and not a distance based directly on allelic differences. Nei's D_A (Nei, 1973), an evolutionary distance, would have been more useful if these were all natural populations and if allele dosage could have been estimated for the AFLP markers. The distance matrices were used to produce dendrograms based on clustering using the unweighted pair-group method with arithmetic averages in the SAHN module of NTSYS-PC program version 2.02 k (Rohlf, 2000). To test the goodness of fit of the clustering to the original data, we used the COPH module to calculate cophenetic value matrices and then compared the original matrix to the cophenetic matrix using MXCOMP.

Bootstrap analysis to validate the dendrogram was difficult to carry out since bootstrap software for SSR-based frequency values does not exist as far as we are aware. However, we performed a principle component analysis (PCA) to validate the dendrograms (Johnson, 1998). Correlation matrices were calculated for marker frequency data for both AFLP and SSR markers, using the similarity/interval module in NTSYS. Ordination was then performed on the correlation matrices to estimate eigenvalues and eigenvectors of each marker type. The original frequency matrices were then ordinated using the eigenvalues and eigenvectors as axes.

Because we were working with an autotetraploid, and previous reports suggested that AFLP and SSR gave the same information, only the SSR data was used to investigate among-population variation. Population differentiation is based on the correlation between alleles taken at random from a sub-population and alleles taken at random from the total population, as measured using Wright's *F* statistics (1951), or related statistics (i.e., Slatkin's [1995] R_{st} and Nei's [1973] G_{st} statistic). If frequencies are the same, F_{st} is zero, and the populations are in panmixia; if frequencies are different, F_{st} approaches 1, and the populations are diverging. If different alleles are fixed in each population, then F_{st} is equal to 1 and the populations are not related. Estimation of population structure using microsatellites can be difficult, considering their complex patterns of mutation, and occurrence of null alleles. Balloux and Lugon-Moulin (2002) recommended using both *F* statistics and *R* statistics to estimate population differentiation. Among the 11 accessions, we estimated Wright's F_{st} (Wright, 1951), using the method of Weir and Cockerham (1984), modified for autotetraploid species using the software, GENE4X (Ronfort et al., 1998). We tested for significance using Fisher's exact tests (Raymond and Rousset, 1995). The software SPAGeDi 1.2 (Hardy and Vekemans, 2002), set for tetraploids, was used. Slatkin's R_{st} (Slatkin, 1995), which is analogous to

F_{st} but assumes a stepwise mutation model, was estimated and tested for significance using permutation tests.

Introgression

We used the software STRUCTURE version 2 (<http://pritch.bsd.uchicago.edu/structure.html>), developed by Pritchard et al. (2000), to examine the occurrence of admixture between the wild and cultivated populations. Their model assumes K populations, each characterized by a set of allele frequencies at each locus. Within populations, loci are assumed to be unlinked, and at linkage equilibrium. Populations are assumed to be in Hardy–Weinberg equilibrium. Based on their genotype, individuals are assigned probabilistically to K clusters, and population allele frequency is simultaneously estimated to find the population grouping that best fits the model (i.e., the grouping most in equilibrium). This is done using a Bayesian approach, using Markov chain Monte Carlo methods. We estimated K using the ad hoc statistic ΔK , which is the second-order rate of change of the estimated log probability (Pr) of data ($\ln \Pr(X|K)$) with respect to the number of clusters, between successive K values divided by the standard error of the probability at K (Pritchard et al., 2000). Evanno et al. (2005) found this value to be best for identifying the appropriate number of clusters in a number of different simulations. Because of the likelihood of common ancestry among the wild collected populations due to geographic proximity and the possibility of allele exchange between the wild populations and Russian cultivars (either inadvertent or during cultivar development), we used the admixture ancestry and correlated allele frequency models developed by Falush et al. (2003), found in Structure, Version 2. We used a burnin of 10,000 and a run length of 20,000 replications. Once we determined K for the 11 accessions based on AFLP and SSR data, we examined the proportion of membership of each accession and compared allele frequency divergence among inferred clusters using the Kullback–Leibler distance, both averaged over individuals, in each inferred cluster.

Marker Congruence

Congruence between the AFLP and SSR markers was assessed by examining the correlation between the pairwise genetic distance matrices generated from both types of markers. Correlation between matrices was examined using the product moment correlations (r) derived from the normalized Mantel Z (Mantel, 1967). Comparisons between matrices were performed using the MXCOMP module of NTSYS-PC program (Rohlf, 2000). The results of the AFLPs and SSRs for measures of diversity were compared using Mann–Whitney’s U-test.

RESULTS AND DISCUSSION

Genetic Diversity

For the SSRs, a total of 118 polymorphic fragments were scored from six loci. Figure 1 shows a representative SSR gel and an example of lanes scored as 1:1:1:1, 1:2:1, 4, and 3:1, respectively. Although deviation from Hardy–Weinberg

equilibrium is more likely to occur in autotetraploid species than in diploid species since more generations are needed to reach equilibrium (Felber and Bever, 1997), Hardy–Weinberg equilibrium had been reached across all populations and across the cultivated and wild germplasm.

Table 3 shows values of the fixation index (F) for each locus and the 11 accessions. Three of the six loci (P001, P003, and P004) showed heterozygote deficiency in 80 to 90% of the accessions. Although polymerase chain reaction products were observed for all individuals, the presence of null alleles may account for heterozygotic deficiency, which was also reported by Flajoulot et al. (2005). When we estimated the frequency of null alleles for the loci showing heterozygotic deficiency, according to Brookfield (1996), mean frequency across accessions was 0.05, 0.18, and 0.2 for P001, P003, and P004, respectively. Given the greater frequency of null alleles, and that heterozygotic deficiency was evident in all but one of the accessions, we dropped P004 from the analysis. After dropping locus P004 from the analysis, we had 111 scorable fragments. There tended to be more unique fragments among the cultivars. For example, ‘Dikoraskaya’ had seven and ‘Semirechenskaya’ had five unique fragments. Among the wild populations, PI 634154 had two unique fragments, and PI 634170 and PI 634145 had one each.

The AFLP analysis resulted in 178 scorable fragments. Seven fragments were found only in the cultivars, and four fragments were found only in the wild populations. PI 634170, collected in a remote and isolated location in the Mūgodžor Mountains, was distinguished by having three fragments not observed in the other 10 accessions.

There was no significant difference between the wild and cultivated accessions for number of alleles (Mann–Whitney U Test; $p = 0.07$) and percentage polymorphic loci (Mann–Whitney U Test; $p = 0.10$). However, gene

Table 3. Values of fixation index (F) for six simple sequence repeat loci in five cultivars and six wild populations of *Medicago sativa*.[†]

Accession	P001	P004	P002	P003	FC16	AW31
PI 634153	0.716***	0.809***	0.454***	0.614***	0.319***	0.337***
PI 634170	0.656***	0.122ns	–0.027ns [‡]	0.207***	0.152***	–0.072ns
PI 634169	0.055ns	0.318**	–0.015ns	0.097ns	–0.062ns	–0.053ns
PI 634145	0.216***	0.288**	–0.076ns	–0.010ns	–0.105ns	–0.093ns
PI 634154	0.615***	0.749***	0.386***	0.543***	0.386***	0.499***
PI 634142	0.655***	0.425*	0.450***	0.794***	0.409***	0.191***
‘Dikorastaskaya’	0.389***	0.819***	0.056ns	0.244***	–0.015ns	0.038**
‘Priaralskaya’	0.176***	0.508***	0.040ns	0.367***	0.104ns	–0.096ns
‘Dolanskaya’	0.129***	0.595***	0.184ns	0.134***	–0.009ns	0.114ns
‘Semirechenskaya’	0.055ns	0.432***	–0.039ns	0.360***	–0.052ns	–0.028ns
‘Tibetskaya’	0.099***	0.406***	–0.042ns	0.254***	0.002ns	–0.009ns

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.

[†]Conformance to Hardy–Weinberg equilibrium was tested using chi-square goodness of fit analysis

[‡]ns, not significant at the 0.05 probability level.

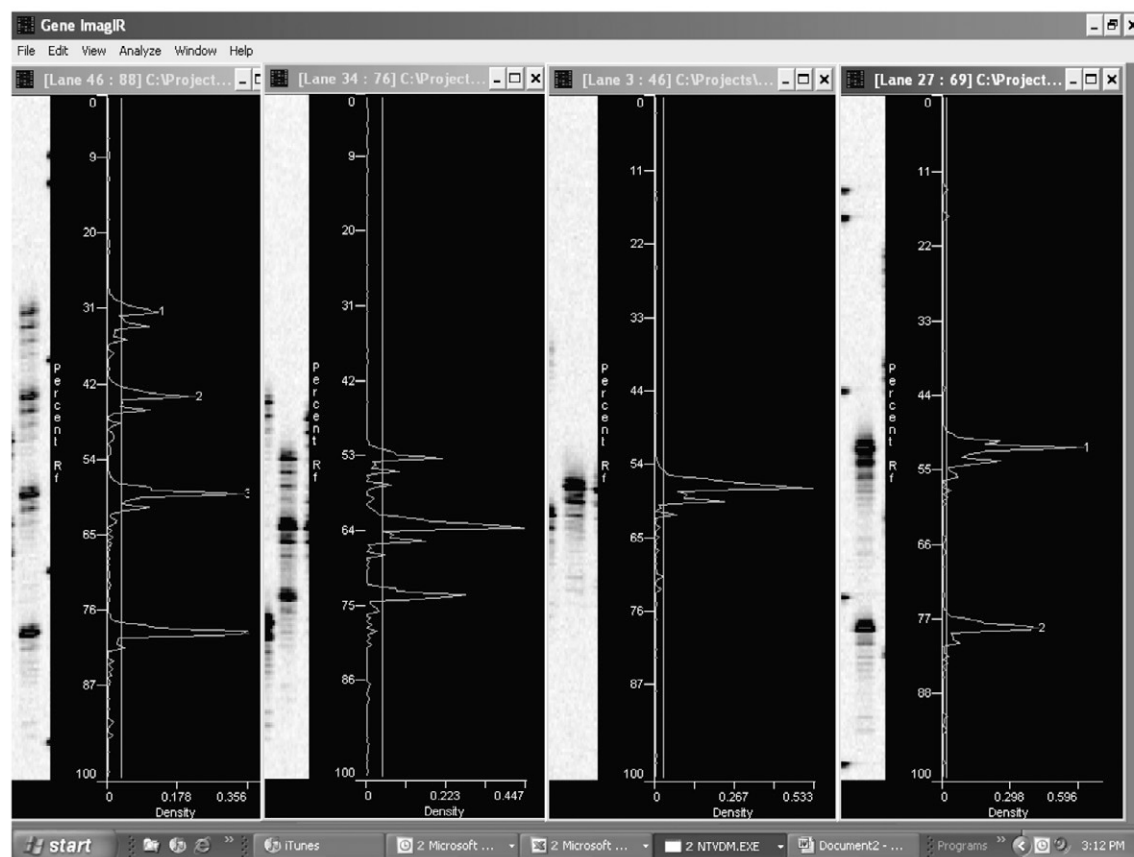
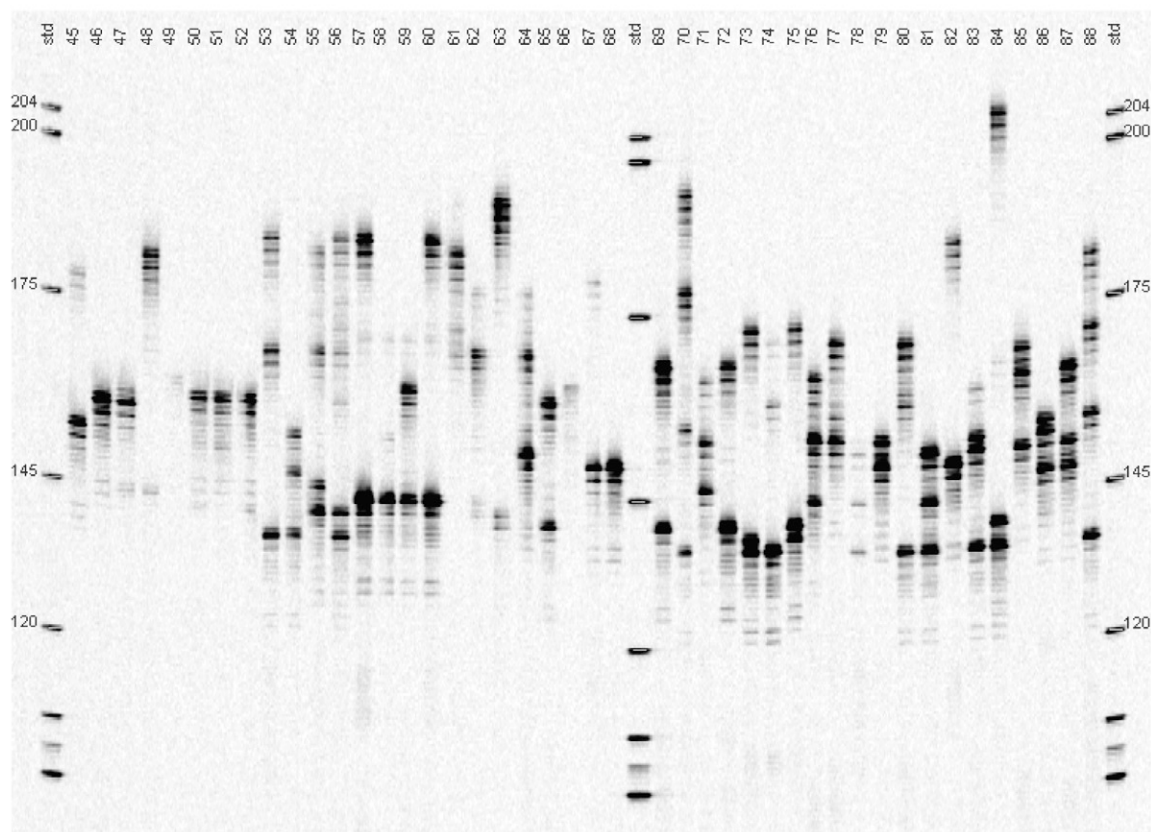


Figure 1. A representative simple sequence repeat gel. Example of lanes that were scored as (left to right) 1:1:1:1, 1:2:1, 4, and 3:1, respectively.

diversity was significantly less in the wild populations compared with the Russian cultivars, measured by SSRs (mean H_E) (Mann–Witney U Test; $p = 0.05$) and AFLPs (mean H) (Mann–Witney U Test; $p = 0.008$), (Table 4).

Genetic Distance and Differentiation

Dendrograms based on AFLP and SSR marker data suggested that the cultivated varieties were more closely related to each other and that the wild populations were more closely related to each other (Fig. 2). The distance matrices and the cophenetic matrices were strongly correlated ($r = 0.85$, $p \leq 0.001$ and $r = 0.98$, $p \leq 0.001$ for SSR and AFLP, respectively) for both markers, indicating that the dendrograms represented the distance matrices with minimal distortion. The PCA analysis gave similar results, further supporting the analysis (data not shown).

The accessions showed significant population differentiation. All pairwise F_{st} estimates were significant ($p < 0.001$; Fisher's exact test) (Table 5.) Genetic differentiation among the cultivars was weak, ranging from 0.019 between 'Tibetskaya' and 'Dolanskaya', to 0.043 between Dolanskaya and 'Priaraskaya'. This was evident in the dendrograms as well (Fig. 2). Populations with the greatest differentiation were PI 634145 with the three wild accessions PI 634169 (0.161), PI 634154 (0.151), and PI 634142 (0.215); and Semirechenskaya with PI 634142 (0.148). The F_{st} values of the remaining populations suggested moderate differentiation (based on interpretation of per se values suggested by Balloux and Lugon-Moulin [2002]). When we estimated differentiation using a stepwise mutation model (Slatkin's R_{st}), only seven pairwise estimates were significant ($p < 0.05$; permutation tests). A drawback of R_{st} is high variance, which may account for lack of significance (Balloux and Lugon-Moulin, 2002). Among the significant values, differentiation was less among the Russian varieties and greatest between the varieties and wild populations. The most differentiated populations were Tibetskaya with PI 634170 (0.202), and PI 634145 (0.297); and Dolanskaya with PI 634154 (0.197) (Table 5).

Introgression

When we used the Bayesian approach and the ad hoc statistic, dK, (Evanno et al., 2005) to delineate the number of clusters based on AFLP markers, the highest value of dK was at $K = 2$ (Fig. 3a). The cultivars contributed an average of 99% of their alleles to cluster 1, while the wild accessions contributed less than 1%. The reverse was true in cluster 2, where the wild accessions contributed 99% of their alleles compared with 1% from the cultivars. A high value of dK at K

$= 5$ suggested further subdivision with limited admixture occurring between the wild and cultivated populations (Table 6). Cluster 1 contained between 91 and 98% of the alleles from the cultivated germplasm and only 0.4 to 1.4% alleles from the wild germplasm. The remaining clusters showed admixture among the wild populations, although in each cluster, one wild population dominated. The exception was PI 634169, whose alleles were distributed among the four clusters (Table 6). Examining allele frequency divergence among the clusters, clusters 1 and 3 (1.17, Kullback–Leibler distance) and clusters 1 and 4 (0.95, Kullback–Leibler distance) were the most divergent. Cluster 1 contained the cultivated germplasm, cluster 3 contained primarily PI 634145, and cluster 4 contained primarily PI 634170. The least-divergent clusters were clusters 2 and 5 (0.12, Kullback–Leibler distance), which both contained a significant percentage of alleles from PI 634153, PI 634169, PI 634154, and PI 634142.

Bayesian analysis of SSR markers using ad hoc dK also revealed a high value at $K = 2$. However, unlike the clear delineation of cultivated and wild accessions, with AFLPs, the SSR markers included a large proportion of alleles from PI 634145, PI 634153, and PI 634154 in the same group as the cultivars. Examination at higher K , however, revealed that this was likely a result of these populations being very distinct, even within the wild populations. The ad hoc statistic dK revealed further subdivisions at $K = 5$ and $K = 8$ (Fig. 3b). When $K = 5$, one cluster was represented primarily by cultivated alleles, and the other four by wild alleles. (Table 7). At $K = 8$, five clusters were represented by each of the five individual wild populations, and three clusters contained an admixture of individuals

Table 4. Summary of variation across all loci in individual populations and means over cultivated and wild groups of *Medicago sativa*, estimated by simple sequence repeat (SSR) and amplified fragment length polymorphism (AFLP) markers.

Accession	N [†]	SSR		AFLP	
		No. of alleles	Mean H_E^{\ddagger}	Mean H^{\ddagger}	% polymorphic loci
PI 634153	16	9.17	0.790	0.18	61.8
PI 634170	16	7.83	0.663	0.19	59.5
PI 634169	16	8.83	0.734	0.17	55.0
PI 634145	16	6.50	0.693	0.09	34.8
PI 634154	16	8.33	0.738	0.18	59.5
PI 634142	16	6.5	0.655	0.19	57.3
WILD	96	7.86	0.712	0.17	54.65
'Dikorastuskaya'	16	10.83	0.812	0.20	58.4
'Priaralskaya'	16	9.17	0.769	0.22	61.8
'Dolanskaya 2'	15	7.50	0.742	0.21	60.1
'Semirechenskaya'	14	10.5	0.770	0.20	61.1
'Tibetskaya'	16	10.0	0.786	0.24	67.4
Cultivated	77	9.6	0.776	0.21	61.76

[†]N, number of individual plants.

[‡]H, heterozygosity; H_E , expected heterozygosity.

from among the cultivated populations. Examining allele frequency divergence at $K = 8$ (data not shown), clusters 5 and 6, which were dominated by PI 645145 and PI 645142, respectively, (2.73, Kullback–Leibler distance) were the most divergent. Clusters 2 and 3 (0.71, Kullback–Leibler distance), which contained mainly cultivated material, were the least divergent.

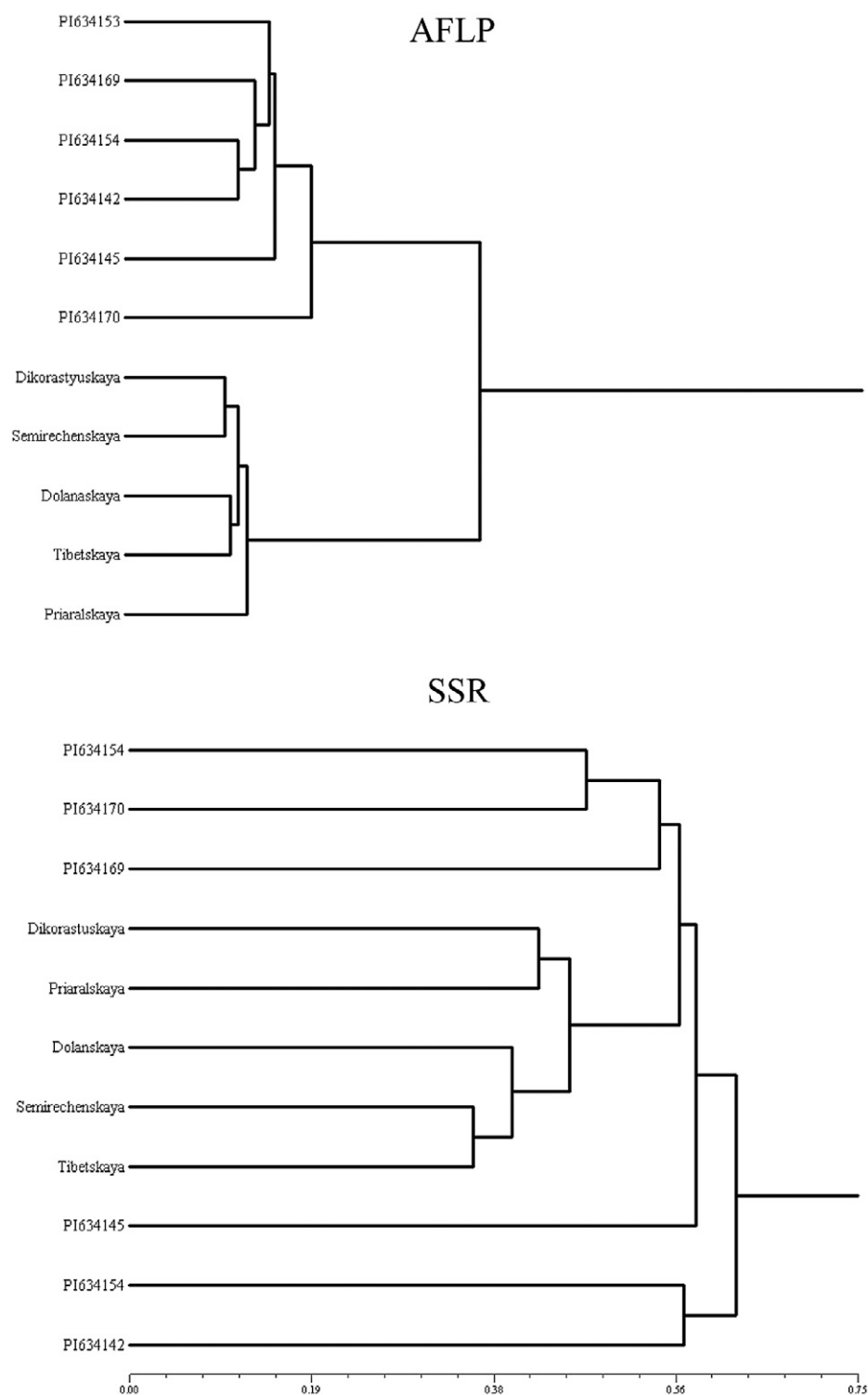


Figure 2. Dendrograms based on amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR) markers, using Provost's distance coefficient. Based on Mantel test, dendrogram correlation was significant but weak ($r = 0.42$, $p \leq 0.01$).

The Bayesian analysis suggested that the wild populations collected in 2000 were minimally contaminated with domesticated alleles. Some admixture was evident, but domesticated alleles generally made up less than 5% admixture in clusters other than clusters composed almost exclusively of the cultivated germplasm. The AFLPs showed limited admixture (5.0%) of Tibetskaya with PI

634153, 634154, 634142, 634169, and 634145. The SSRs also showed limited admixture (5.8%) of Tibetskaya with PI 634153, 634154, 634142, 634145, and 634170. In examining the pedigree of Tibetskaya, it was found to have originated from a Vavilov Institute of Plant Industry germplasm accession collected in Tibet, which was subsequently developed into a cultivar during the 1930s at the Kuban Experimental Station, in the Krasnodar region of the North Caucasus Mountains. However, extensive breeding work on this germplasm was also performed at the Priaral Experiment Station, which is located in the region of our 2000 collection trip. Admixture of the wild populations with Semirechenskaya was also evident, but low ($<5\%$). Semirechenskaya was developed from a landrace grown in the south of Kazakhstan (Alma-Ata region), and recognized as a variety in 1934. This cultivar would have been used in breeding all the studied varieties except Tibetskaya. A small amount of admixture was also evident between the wild populations and 'Dikorastuskaya' (1.4%). Considering that the Russian word *Dikorastuskaya* means "wild growing," the accession may actually be wild-growing material (as opposed to a cultivated form) collected in the Aktobe region before the 1950s. The close relationship of Dikorastuskaya with the other cultivated forms may reflect a snapshot of the introgression that occurred with Russian varieties grown during that time period. The low level of admixture between 'Priaralskaya' and the wild populations was surprising. This variety was developed at the Priaral Experiment Station, during A.I. Ivanov's tenure, using germplasm collected in the Mugodzor Mountains (N.I. Dzyubenko, personal communication, 2006), which was an area collected from in 2000.

The low incidence of admixture between the Russian varieties and wild populations may reflect the unique adaptation of the wild populations to the dry, cold, and largely ungrazed conditions prevalent in the region during the 2000 visit. Similar to studies reported on *M. sativa* ssp. *sativa* Mielga in southern Spain (Jenczewski et al., 1999a,b; Muller et al., 2001); wild-type alleles may convey better adaptation to the harsh environment than domesticated alleles. However, there was some evidence that introgression occurred in the past. PI 634153 and PI 634154 had the most amount of admixture with the cultivated germplasm, having 8 and 11.6% membership in SSR cluster 2, which contained primarily Russian varieties. These accessions were collected in the vicinity of the Bol'shaya Khobda River, an area observed to be the most populated and having a long history of grazing and forage production in the 2000 survey (Greene et al., 2005). In contrast, PI 634169 and PI 634170 had the most unique allele profiles and the least amount of admixture with the other populations. These two accessions were collected in the remote southern Mūgodžor Mountains, an area that has had minimal cultivation.

Marker Congruence

Regarding genetic variation between the wild and cultivated germplasm, the SSR and AFLP markers gave the same results. Both markers indicated that there was no difference between wild and cultivated forms for number of alleles and percentage polymorphic loci but that diversity was less among and within the wild forms than in cultivated forms. The similarity between the two dendrograms is evident (Fig. 2). Both markers

clustered the cultivated accessions together, and among the wild populations, both markers grouped PI 634154 and PI 634142 together, apart from the other four wild accessions. When we compared the genetic distance matrices generated by the AFLPs and SSRs, using a Mantel test, the correlation was significant but weak ($r = 0.42$, $p \leq 0.01$).

Table 5. Pairwise F_{st} (top half of matrix) and R_{st} (bottom half of matrix) values for wild populations and Russian varieties of *Medicago sativa*, calculated with five simple sequence repeat loci.[†]

Accession	PI 634153	PI 634170	PI 634169	PI 634145	PI 634154	PI 634142	'Dikora-staskaya'	'Priaral-skaya'	'Dolan-skaya'	'Semirech-enskaya'	'Tibet-skaya'
PI 634153		0.037	0.087	0.122	0.058	0.078	0.024	0.058	0.051	0.077	0.053
PI 634170	0.066		0.081	0.135	0.083	0.077	0.063	0.079	0.080	0.106	0.086
PI 634169	0.169	0.202		0.161	0.091	0.121	0.089	0.105	0.113	0.115	0.084
PI 634145	0.232	0.316	0.425		0.151	0.215	0.103	0.104	0.112	0.087	0.108
PI 634154	0.087	0.145	0.178	0.283		0.102	0.065	0.085	0.108	0.107	0.077
PI 634142	0.11	0.133	0.227	0.375	0.148		0.099	0.113	0.139	0.148	0.133
'Dikorastaskaya'	0.045	0.141	0.224	0.261	0.120	0.175		0.029	0.033	0.035	0.034
'Priaralskaya'	0.107	0.179	0.268	0.274	0.156*	0.201	0.073		0.043	0.026	0.042
'Dolanskaya'	0.097	0.182	0.290	0.295	0.197*	0.242	0.085	0.112		0.023	0.019
'Semirechenskaya'	0.143	0.240	0.305	0.251	0.198	0.259	0.092	0.072	0.066		0.024
'Tibetskaya'	0.104	0.202*	0.236	0.297*	0.150*	0.242	0.089	0.113*	0.055*	0.071	

* R_{st} values significant at the 0.05 probability level; alleles permutation tests.

[†]All F_{st} values significant at the 0.001 probability level; Fisher's exact test.

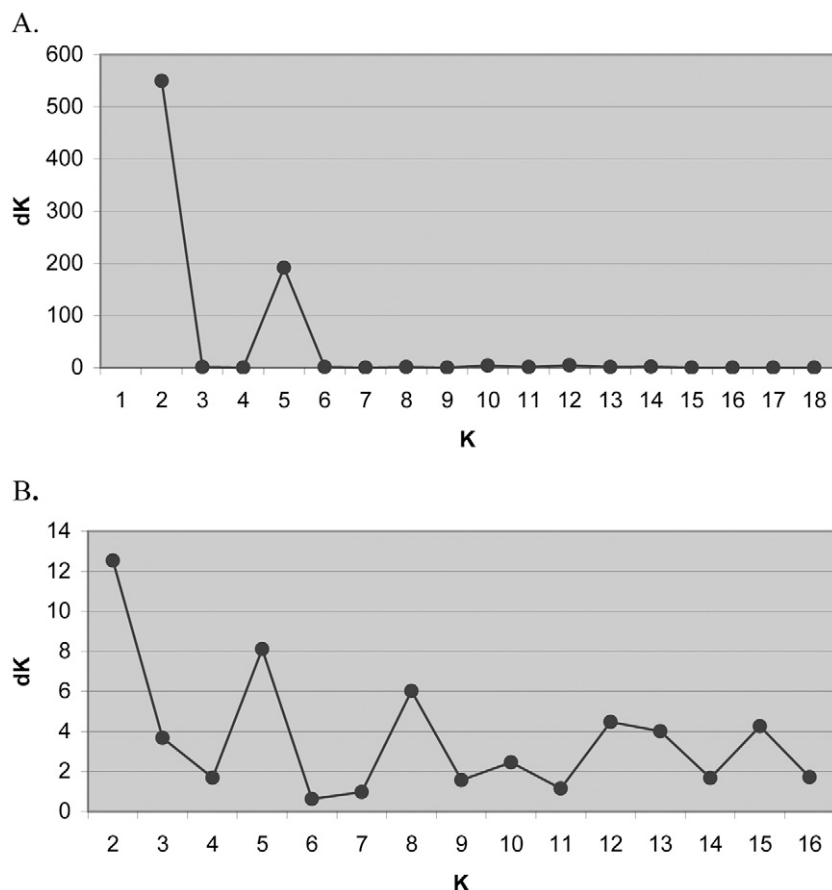


Figure 3. The statistic dK indicates the estimated number of clusters based on amplified fragment length polymorphism markers is five. The estimated number of clusters based on simple sequence repeat markers is also five, although there is some indication of further subdivision at $K = 8$, where K is a group of populations characterized by a similar set of allele frequencies at each locus.

Table 6. Percentage of each *Medicago sativa* population's alleles that fall into $K = 5$ clusters using amplified fragment length polymorphism marker data (K = group of populations characterized by a similar set of allele frequencies at each locus).

Accession	1	2	3	4	5
	%				
PI 634153	0.9	79.5	3.1	2.5	13.9
PI 634170	0.7	1.3	2.1	89.0	6.9
PI 634169	0.5	31.3	14.4	14.2	39.5
PI 634145	0.7	1.4	93.7	1.6	2.6
PI 634154	0.4	37.7	1.4	1.8	58.7
PI 634142	1.4	21.4	2.7	1.8	72.7
'Dikorastaskaya'	96.9	0.6	0.7	0.4	1.4
'Priaralskaya'	98.4	0.4	0.5	0.3	0.4
'Dolanskaya'	98.4	0.4	0.4	0.3	0.4
'Semirechenskaya'	98.4	0.3	0.6	0.3	0.4
'Tibetskaya'	91.3	1.6	1.7	0.5	4.8

Table 7. Percentage of each *Medicago sativa* population's alleles that fall into $K = 5$ clusters using simple sequence repeat marker data (K = group of populations characterized by a similar set of allele frequencies at each locus).

Accession	1	2	3	4	5
	%				
PI 634153	13.7	13.2	13.9	27.5	31.8
PI 634170	1.6	37.2	10.8	3.5	46.8
PI 634169	2.8	89.6	3.3	1.5	2.9
PI 634145	1.1	2.9	88.8	2.9	4.3
PI 634154	9.2	5.9	1.9	60.5	22.5
PI 634142	0.9	44.0	1.0	43.3	10.7
'Dikorastaskaya'	48.9	3.6	14.9	14.4	18.3
'Priaralskaya'	57.0	9.0	5.3	7.9	20.8
'Dolanskaya'	82.0	1.0	1.8	6.0	9.2
'Semirechenskaya'	74.7	7.4	6.4	3.6	7.9
'Tibetskaya'	79.0	2.9	6.6	7.9	3.7

Based on the cluster analysis, accessions were more closely related when AFLPs were used compared with the SSR markers. Values for allele-frequency divergence among K clusters based on AFLP markers also tended to be less than values estimated by the SSR markers. Generally, separation between the cultivated and wild accessions was not as clear with the SSRs as the AFLPs. Similar results have been reported elsewhere and have been attributed to the hypervariability and higher resolving power of SSR loci (Powell et al., 1996).

Although the AFLP and SSR markers provided congruent information, using both markers provided more information than if we had used either marker alone. The AFLPs confirmed the separation between the wild and cultivated accessions, which was not as evident from the SSR data. Gaudeul et al. (2004) reported that AFLPs could generally distinguish between close populations, whereas SSR markers could not. However, the SSRs were effective

for understanding the fine-scale structure within the wild populations. Gaudeul et al. (2004) also reported that SSRs were better for resolving fine-scale spatial structure within populations and that both markers were generally congruent when there were large differences in populations.

CONCLUSIONS

Hybridization between introduced and wild plant populations is a common occurrence in areas where both forms grow sympatrically. Considering the close relationship with *Dikorastaskaya*, a "wild-growing" accession collected in the 1950s, and the other four Russian cultivars, and higher levels of admixture among populations collected in more populous areas, our study suggests that introgression has occurred in the past but that it was likely a result of human intervention as wild populations were collected for breeding purposes. Although in many instances, introgression results in the genetic erosion of wild populations, our study suggests that the consequences of gene flow between Russian alfalfa varieties historically grown in northwestern Kazakhstan and native population of wild relatives have been minimal considering the small amount of admixture we observed between wild populations collected in 2000 and Russian cultivars. Similar to the Spanish ecotype Mielga (Jenczewski et al., 1999a,b; Muller et al., 2001, 2003), domestic alleles may be less adapted to the wild environment and not persistent over time. We found not only that wild populations collected in 2000 had limited contamination with Russian varieties but that some of the wild populations stood out as valuable genetic resources compared with the other wild populations we studied. PI 634169 and PI 634170 collected in the remote Southern Mūgodžor Mountains had unique allele profiles and the least amount of admixture with the other populations. In the 2000 survey, we noted that the Southern Mūgodžor Mountains had a rich diversity of species, suggesting that this area should be further assessed to determine the feasibility of establishing in situ reserves (Greene et al., 2005). PI 634145 also had a unique profile, having 93.5% membership in a single cluster based on AFLPs and 88% membership in a single cluster based on SSRs. This accession was collected in the northern grassland steppes in an area where diploid and tetraploid forms of *M. sativa* subsp. *falcata*, diploid *M. sativa* subsp. *coerulea*, tetraploid *M. sativa* subsp. *sativa*, and their respective diploid (*M. sativa* nothosubsp. *hemicycla*) and tetraploid (*M. sativa* nothosubsp. *varia*) hybrid forms were collected. This area would also benefit from in situ protection. Finally, we found that using both AFLP and SSR markers allowed us to better understand the population structure and consequences of introgression that have occurred between wild alfalfa relatives and alfalfa cultivars historically grown in northwestern Kazakhstan.

Acknowledgments

S. Greene and N. Dzubenko wish to thank the other members of the 2000 Kazakhstan exploration trip, R. Hannan, A. Afonin, and A. Khusainov and acknowledge the funding provided by the USDA National Plant Germplasm System to support the collecting trip.

References

- Balloux, F., and N. Lugon-Moulin. 2002. The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* 11:155–165.
- Boudry, P., M. Morchen, P. Saumitou-Laprade, P. Vernet, and H. Van Dijk. 1993. The origin and evolution of weed beets: Consequences for the breeding and release of herbicide-resistant transgenic sugar beets. *Theor. Appl. Genet.* 87:471–478.
- Brookfield, J.F.Y. 1996. A simple new method for estimating null allele frequency from heterozygous deficiency. *Mol. Ecol.* 5:453–455.
- de Wet, J.M., and J.R. Harlan. 1975. Weeds and domesticates: Evolution in the man-made habitat. *Econ. Bot.* 29:99–107.
- Diwan, N., A.A. Bhagwat, G.R. Baughan, P.B. Cregan. 1997. Simple sequence repeat (SSR) DNA markers in alfalfa and perennial and annual *Medicago* species. *Genome* 40:887–895.
- Ellstrand, N.C. 2003. Dangerous liaisons? When cultivated plants mate with their wild relatives. John Hopkins Univ. Press, Baltimore, MD.
- Ellstrand, N.C., H.C. Prentice, and J.F. Hancock. 1999. Gene flow and introgression from domesticated plants into their wild relatives. *Annu. Rev. Ecosyst.* 30:539–563.
- Ellwood, S.R., N.K. D'Souza, L.G. Kamphuis, T.I. Burgess, R.M. Nair, and R.P. Oliver. 2006. SSR analysis of the *Medicago truncatula* SARDI core collection reveals substantial diversity and unusual genotype dispersal throughout the Mediterranean basin. *Theor. Appl. Genet.* 112:977–983.
- Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14:2611–2620.
- Falush, D., M. Stephens, and J.K. Pritchard. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164:1567–1587.
- Felber, F., and J.D. Bever. 1997. Effect of triploid fitness on the coexistence of diploids and tetraploids. *Biol. J. Linn. Soc.* 60:95–106.
- Flajoulot, S., J. Ronfort, P. Baudouin, P. Barre, T. Huguet, C. Huyghe, and B. Julier. 2005. Genetic diversity among alfalfa (*Medicago sativa*) cultivars coming from a breeding program, using SSR markers. *Theor. Appl. Genet.* 111:1420–1429.
- Gaudeul, M., I. Till-Bottraud, F. Barjon, and S. Manel. 2004. Genetic diversity and differentiation in *Eryngium alpinum* L. (Apiaceae): Comparison of AFLP and microsatellite markers. *Heredity* 92:508–518.
- Greene, S.L., R. Hannan, A. Afonin, N.I. Dzyubenko, and A. Khusainov. 2005. Collecting wild crop relatives in the northwestern steppes of Kazakhstan. *Plant Genet. Res. Newsl.* 141:1–7.
- Hails, R.S., and K. Morley. 2005. Genes invading new populations: A risk assessment perspective. *Trends Ecol. Evol.* 20:245–252.
- Hardy, O.J., and X. Vekemans. 2002. SPAGeDi: A versatile computer program to analyze spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* 2:618–620.
- Hoc, P.S., S.M. Espert, S.I. Drewes, and A.D. Burghardt. 2006. Hybridization between wild and domesticated types of *Phaseolus vulgaris* L. (Fabaceae) in Argentina. *Genet. Res. Crop Evol.* 53:331–337.
- Ivanov, A.I. 1988. Alfalfa. Amerind Publishing, New Delhi, India.
- Jarvis, D.I., and T. Hodgkin. 1999. Wild relatives and crop cultivars: Detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. *Mol. Ecol.* 8:S159–S173.
- Jenczewski, E., J.M. Prosperi, and J. Ronfort. 1999a. Differentiation between natural and cultivated populations of *Medicago sativa* (Leguminosae) from Spain: Analysis with random amplified polymorphic DNA (RAPD) markers and comparison with allozymes. *Mol. Ecol.* 8:1317–1330.
- Jenczewski, E., J.M. Prosperi, and J. Ronfort. 1999b. Evidence for gene flow between wild and cultivated *Medicago sativa* (Leguminosae) based on allozyme markers and quantitative traits. *Am. J. Bot.* 86:677–687.
- Johnson, D.E. 1998. Applied multivariate methods for data analysts. Brooks-Cole Publishing, Monterey, CA.
- Julier, B., S. Flajoulot, P. Barre, G. Cardinet, S. Santoni, T. Huguet, and C. Huyghe. 2003. Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. *BMC Plant Biol.* 3:9.
- Lefèvre, F. 2004. Human impacts on forest genetic resources in the temperate zone: An updated review. *For. Ecol. Manage.* 197:257–271.
- Lesins, K.A., and I. Lesins. 1979. Genus *Medicago* (Leguminosae). A taxogenetic study. Dr. W. Junk BV Publishers, The Hague.
- Mantel, N.A. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27:209–220.
- Mariette, S., V. Le Corre, F. Austerlitz, and A. Kremer. 2002. Sampling within the genome for measuring within population diversity: Trade-offs between markers. *Mol. Ecol.* 11:1145–1156.
- Meilleur, B.A., and T. Hodgkin. 2004. In situ conservation of crop wild relatives: Status and trends. *Biodivers. Cons.* 13:663–684.
- Michaud, R., W.F. Lehman, and M.D. Rumbaugh. 1988. World distribution and historical development. p. 25–91. In A.A. Hanson, D.K. Barnes, and R.R. Hill (ed.) Alfalfa and alfalfa improvement. Agron. Monogr. 29. ASA, CSSA, and SSSA, Madison, WI.
- Muller, M.H., J.M. Prosperi, S. Santoni, and J. Ronfort. 2001. How mitochondrial DNA diversity can help to understand the dynamics of wild-cultivated complexes. The case of *Medicago sativa* in Spain. *Mol. Ecol.* 10:2753–2763.
- Muller, M.H., J.M. Prosperi, S. Santoni, and J. Ronfort. 2003. Inferences from mitochondrial DNA patterns on domestication history of alfalfa (*Medicago sativa*). *Mol. Ecol.* 12:2187–2199.
- Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* 70:3321–3323.
- Nybom, H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Mol. Ecol.* 13:1143–1155.
- Oka, H.I. 1992. Ecology of wild rice planted in Taiwan: II. Comparison of two populations with different genotypes. *Bot. Bull. Acad. Sinica (Taiwan)* 33:75–84.
- Panetsos, C.A., and H.G. Baker. 1967. The origin of variation in “wild” *Raphanus sativus* (Cruciferae) in California. *Genetica* 38:243–274.
- Powell, W., M. Morgante, C. Andre, M. Hanafey, and J. Vogel. 1996. The comparison of RFLP, RAPD, AFLP, and SSR

- (microsatellite) markers for germplasm analysis. *Mol. Breed.* 2:225–238.
- Prevosti, A., J. Ocana, and G. Alonso. 1975. Distances between populations of *Drosophila subobscura*, based on chromosome rearrangement frequencies. *Theor. Appl. Genet.* 45:231–241.
- Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Raymond, M., and F. Rousset. 1995. An exact test for population differentiation. *Evolution Int. J. Org. Evolution* 49:1280–1283.
- Rohlf, F.J. 2000. Numeric taxonomy and multivariate analysis system (NTSYSpc), version 2.1. Exeter Software, New York.
- Ronfort, J.E., J. Jenczewski, T. Bataillon, and F. Rousset. 1998. Analysis of population structure in autotetraploid species. *Genetics* 150:921–930.
- Rufener Al Mazyad, P., and K. Ammann. 1999. Biogeographical assay and natural gene flow. In K. Ammann, Y. Jacot, V. Simonsen, and G. Kjelsson (ed.) *Methods for risk assessment of transgenic plants: III. Ecological risks and prospects of transgenic plants*. Birkhauser, Basel, Switzerland.
- Sinskaya, E.N. 1961. Flora of the cultivated plants of the USSR. Vol. 13. Perennial leguminous plants. National Science Foundation, Washington, DC.
- Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139:457–462.
- Song, Z., B. Li, J. Chen, and B. Lu. 2005. Genetic diversity and conservation of common wild rice (*Oryza rufipogon*) in China. *Plant Species Biol.* 20:83–92.
- Thrall, P.H., and A. Young. 2000. Autotet: A program for analysis of autotetraploid genotypic data. *J. Hered.* 91:348–349.
- Weir, B.S., and C.C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. *Evolution Int. J. Org. Evolution* 38:1358–1370.
- Wiersema, J.H., K.J. Kirkbride, and C.R. Gunn. 1990. Legume (Fabaceae) nomenclature in the USDA germplasm system. USDA Technical Bull. No. 1757. U.S. Gov. Print. Office, Washington, DC.
- Woodhead, M., J. Rusell, J. Squirell, P.M. Hollingsworth, K. Mackenzie, M. Gibby, and W. Powell. 2005. Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. *Mol. Ecol.* 14:1681–1695.
- Wright, S. 1951. The genetical structure of populations. *Ann. Eugen.* 15:323–354.
- Yeh, F.C., R. Yang, and T. Boyle. 1999. POPGENE: Microsoft Windows-based freeware for population genetic analysis. Release 1.31. University of Alberta, Edmonton, AB, Canada.